PhD proposal: Super-resolved generative and space-time adaptive neural operator for 3D extreme-scale wave propagation problems.

Abstract

This PhD thesis proposal addresses the limitations of numerical simulations in extreme-scale engineering modeling, particularly regarding uncertainty quantification and high computational costs. Neural operators will be investigated to design effective surrogate models that accelerate evaluations of partial differential equations, particularly in the context of broadband regional seismic wave propagation.

Key objectives of the PhD thesis include scaling neural operators for real-world 3D problems, overcoming spatial domain size limitations, and leveraging multiscale approaches. The proposal emphasizes developing transfer learning strategies to handle dataset shifts and enhance model efficiency. Additionally, it aims to investigate data augmentation techniques and super-resolution generative methods for realistic broadband solutions. Finally, a framework based on neural operators will be developed for uncertainty quantification purpose, focusing on posterior probability distributions in high-dimensional spaces through diffusion models. This combination of generative AI and neural operators will be tested on real case scenarios, involving probabilistic seismic hazard estimation over large regions with poor historical observations.

Context

Numerical simulations are key elements in extreme-scale engineering modeling. However, their applicability becomes limited in the context of uncertainty due to their high computational costs. Moreover, due to numerical discretization, the outcome of such simulations is systematically polluted by high frequency noise. Such noise, combined with the lack of direct characterization of the sub-wavelength features, hampers the robust broad-band uncertainty quantification. In recent years, neural operators have emerged as accurate surrogate models to replace numerical solvers by providing fast evaluations.

As far as Scientific Machine learning (SciML) is concerned, several research streams can be identified. We have shown that medium-scale 3D transient elastodynamic problems can be solved using Neural Operator at reasonable costs thanks to transfer learning and data augmentation [1]. Yet this approach has not reached the required versatility and fine spatial and temporal resolution that would be required by practical applications, for instance in risk assessment. Moreover, it is strongly based on synthetic data from high performance numerical simulations, with a very limited use of existing scarce but numerous experimental data.

Another research stream involves diffusion models that can reproduce high-frequency features, even in 3D problems [2]. Due to their intrinsic probabilistic framework, diffusion models are promising candidates to quantify uncertainties related to unresolved small scales. Because of the numerical dispersion at high frequency, synthetic output is generally low-pass filtered. This implies that the target distributions must be obtained from earthquake records. SciML approaches are then envisioned as a bridge between simulation and real data to augment existing approaches [3].

At last, both approaches are unexplainable black-boxes for which the underlying physics is embedded in the training dataset, as opposed to physics-informed approaches which account for the field equations in the architecture. It is believed that invariance properties of neural operators with respect to both data augmentation and transfer learning might bring some level of explainability to these black-box models, together with some robustness and trustworthiness.

Objectives

This thesis aims to propose innovative approaches to scale neural operators to real-world 3D problems. One first challenge is related to the spatial size of the domain in practical applications, which covers a few tens of kilometres with a spatial resolution of tens of meters. Most approaches rely on regular grids that cannot accommodate these dimensions in 3D [4]. However, in engineering practice, high spatial resolution is needed only at given regions. Multiscale neural operators then appear as a promising solution [5].

The second challenge comes from the difficulty of obtaining training data, which involves running computationally demanding numerical solvers. A large database of 30,000 training samples is available [6]. This thesis will develop transfer learning strategies to expand neural operators under strong dataset shifts. This involves, for instance, encoding information from a distant earthquake source that does not belong to the training domain. As such, this work promotes data frugality and reasonable resource usage, without compromising on accuracy. Data augmentation leveraging the geometric structure of the data manifold will especially be investigated in this direction.

Emphasis will be put on understanding the conditions that make transfer learning successful. In particular, the relationship between parameter update and task shift will be quantified to provide more physical understanding of the methods.

Moreover, super-resolution generative strategies will be introduced to provide realistic broadband solutions, conditioned by the entity of the phenomenon (e.g., earthquake magnitude, source type etc.). These high-frequency solutions should maintain the physical features of numerical simulations while matching the temporal characteristics of recordings.

The neural operator will then be deployed for uncertainty quantification analyses in a variational framework. This requires properly defining the probability distributions of uncertain variables, which is a significant challenge in high-dimensional spaces. Diffusion models will be explored to tackle this task.

The main expected outcomes of the thesis are:

- Implement a multiscale 3D neural operator for large spatial domains and high resolutions, and apply it on real case scenarios
- Propose an analysis framework for transfer learning between different datasets of wave propagation, with the aim of being as general as possible
- Integrate the neural operator within a probabilistic framework to allow high-frequency predictions and uncertainty quantification

Publications in the top international machine learning conferences and highly-ranked scientific journals are anticipated. Code and benchmarks produced during the PhD will be released open access.

Supervision

The PhD will be a joint supervision between the Laboratoire de Mécanique Paris-Saclay (Université Paris-Saclay, France) and the Computational and Applied Mathematics Laboratory (ETH Zürich, Switzerland). The PhD candidate will benefit from a rich and stimulating environment with seminars, workshops, informal gatherings, and weekly formal meetings. Up to three months can be spent at ETH Zürich, during which the candidate will have the opportunity to deploy their methodology on different scientific problems.

This PhD topic is participating to the Université Paris-Saclay EU COFUND: DeMythif.AI program. It is especially advertised to international students who have spent less than 12 months in France in the last 3 years. The candidates will be evaluated by a jury who will select 20 PhD to start in fall 2025. The successful candidates will be fully funded for 3 years (monthly gross salary 2 430 €, mobility indemnities 3 000€ and financial support to the host laboratories 10 000 €), have access to specific scientific and non-scientific training, and be fully part of the Université Paris-Saclay AI community.

Supervision: Prof. Didier Clouteau didier.clouteau@centralesupelec.fr and **Dr. Filippo Gatti** filippo.gatti@centralesupelec.fr

Co-supervision: Prof. Siddhartha Mishra and **Dr. Fanny Lehmann**

Requirements

The candidate must hold a Masters degree in Computer Science, Machine Learning, or suitable related field. The candidate should have strong background in programming (Python) and machine learning. Interest in physics and/or applied mathematics will be appreciated.

The candidate must not have spent more than 12 months in France within the last three years.

The deadline for application is January 17, 2025. More information and application on Adum and DataIA PhD call.

References

[1] Fanny Lehmann, Filippo Gatti, and Didier Clouteau. Multiple-Input Fourier Neural Operator (MIFNO) for source-dependent 3D elastodynamics, 2024.

[2] Roberto Molinaro, Samuel Lanthaler, Bogdan Raonić, Tobias Rohner, Victor Armegioiu, Zhong Yi Wan, Fei Sha, Siddhartha Mishra, and Leonardo Zepeda-Núñez. Generative AI for fast and accurate Statistical Computation of Fluids, 2024.

[3] Filippo Gatti and Didier Clouteau. Towards blending Physics-Based numerical simulations and seismic databases using Generative Adversarial Network. *Computer Methods in Applied Mechanics and Engineering*, 372:113421, December 2020.

[4] Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier Neural Operator for Parametric Partial Differential Equations. In *International Conference on Learning Representations*, 2021.

[5] Gege Wen, Zongyi Li, Qirui Long, Kamyar Azizzadenesheli, Anima Anandkumar, and Sally M. Benson. Real-time high-resolution CO2 geological storage prediction using nested Fourier neural operators. *Energy & Environmental Science*, 16(4):1732–1741, April 2023.

[6] F. Lehmann, F. Gatti, M. Bertin, and D. Clouteau. Synthetic ground motions in heterogeneous geologies from various sources: The HEMEW^{S}-3D database. *Earth System Science Data*, 16(9):3949–3972, 2024.